

REMARKS

Claims 1-20 remain pending in the application without amendment.

Based on the following Remarks, Applicant respectfully requests that the examiner reconsider all outstanding objections and rejections, and withdraw them.

The February 12, 2004 Office Action

In the Office Action dated February 12, 2004, the following rejections were asserted:

- ¶ 2-8. Claims 1-2(11-12) and 7(17) were rejected under 35 U.S.C. § 102 as being anticipated by U.S. Patent No. 5,799,002 (Krishnan).
- ¶ 9-16. Claims 3-5 and 13-15 were rejected under 35 U.S.C. § 103 as being unpatentable over Krishnan in view of U.S. Patent No. 6,154,769 (Cherkasova *et al.*).
- ¶ 17-25. Claims 6(16) and 8-10(18-20) were rejected under 35 U.S.C. § 103 as being unpatentable over Krishnan in view of U.S. Patent No. 6,510,214 (McDaniel).

Rejection under 35 U.S.C. § 102(b)

Claims 1(11), 2(12), and 7(17) were rejected under 35 U.S.C. § 102 as being anticipated by U.S. Patent No. 5,799,002 (Krishnan).

Applicants respectfully traverse the rejection.

Preliminarily, Applicants explain that the claims are directed to methods and apparatus for throttling incoming requests to a server that may be, or may not be, *overloaded*. If the server is overloaded, then incoming requests are stored in a buffer *only if* a number of incoming requests in the buffer is less than an acceptance limit. If the server is *not* overloaded, then the incoming requests are stored in the buffer. Claim 1, being indicative of Claim 11, recites:

1. (Original) A method for throttling incoming requests to a server having a buffer for storing incoming calls request to being processed by the server, the method comprising:

receiving incoming requests;
determining whether the **server** is **overloaded**;
if the **server** is **overloaded** then storing the incoming requests in the buffer only if a number of incoming requests currently stored in the buffer is less than an acceptance limit; and
if the **server** is not **overloaded**, the storing the incoming request in the buffer. (emphasis added)

In the exemplary embodiment of the invention to which the claims should not be limited (see top of Applicants' page 4 and Applicants' FIG. 1), a throttling unit 114 (detailed in FIG. 2) controls the number of incoming requests that are transmitted to a server 116 (detailed in FIG. 3). Server 116 includes a buffer 302 (FIG. 3) that is configured to store requests 204. *When the server is overloaded* (a condition positively recited in Claim 1), if the number of requests currently stored in buffer 302 is less than "acceptance limit" 306 (FIG. 3), then incoming requests are stored in the buffer; but if the *server* is not overloaded, then incoming requests are stored in the buffer without regard to the acceptance limit.

In short, the important decision governing operation of the claimed invention relates to whether or not the *server* (example: Applicants' element 116 in FIGS. 1 and 3) is overloaded. In distinct contrast to Applicants' Claim 1, the Krishnan patent discloses a system that focuses on throttling requests based on *network bandwidth* of network N (Krishnan's FIG. 1) and not based on the "overload"-status of any network server NS_n. Applicants respectfully submit that because of this fundamental difference, the Krishnan patent does not properly disclose or suggest the present claimed invention.

Indeed, Krishnan is concerned with minimizing network congestion and utilizing network bandwidth in network N; the Krishnan patent does not disclose or suggest monitoring whether or not a *server* NS_n is overloaded as required by Applicants' claims. In Applicants' claims, it is assumed that the network's capacity is irrelevant but that the *server* may be overloaded; in contrast, in Krishnan's scenario, it is assumed that the servers have unlimited capacity but that the *network* may be congested ("overloaded"). Thus, in relevant part, Krishnan's *network*-centric system is actually the *opposite* of Applicants' *server*-centric claims.

Applicants now discuss Krishnan's disclosure in more detail, so that the examiner may more easily appreciate how the Krishnan patent does not render Applicants' claims unpatentable.

First, Krishnan's "PROBLEM" section recites:

....The network typically has a predefined **bandwidth capacity** which is shared among a number of interconnected computer systems. Each computer system is typically allotted a limited fixed allocation of **network bandwidth** through which to serve requests received from the client processes. When the allocated network bandwidth becomes saturated with client requests and server responses, some of the data is either delayed in transmission or not delivered to the intended destination. Therefore, some form of

request throttling mechanism is necessary to **minimize network congestion and efficiently utilize the allocated network bandwidth.**

In the case of multiple network servers executing on a single host computer system and sharing a **fixed bandwidth communication link to the network**, some network servers can disproportionately allocate this network bandwidth to their tasks, thereby excluding other concurrently executing network servers from performing their requested operations. In this case, the bandwidth throttling must be effected among the plurality of network servers which are concurrently executing on the host computer system.

It is therefore a problem to **allocate bandwidth** to the network server processes in a manner which enables the maximum number of requests to be served without network congestion and to also avoid impacting other network servers which may be executing on the same host computer system. (column 1, lines 17-44; emphasis added)

In contrast to Krishnan's network bandwidth optimization among plural network servers NSn, the method of Applicants' Claim 1 optimizes input of requests to a *server* based on whether the *server* is busy. Thus, even if Krishnan's arrangement optimizes *network* bandwidth allocation, it does not perform the required task of optimizing input of requests to any given *server*, especially in the manner recited in Claim 1 (based on whether the server is overloaded).

The Office Action refers specifically to the following passage in the Krishnan patent, within Krishnan's "SOLUTION" section:

...the present invention which provides a hierarchically organized response to **network congestion** to escalate the actions taken to mitigate the traffic presented to the network **in response to various levels of congestion.** The bandwidth throttling system operates on a host computer system to **allocate bandwidth** to the network servers which are executing on the host computer system as a function of system administrator defined thresholds. (column 2, lines 48-55; emphasis added)

Again: the fact that Krishnan performs his control function based on levels of network congestion, rather than on a *server's* overload status, demonstrates that the Krishnan patent has little relevance to the claimed invention.

Concerning Claim 1's recitation of "determining whether the **server** is **overloaded**," the Office Action specifically refers to the following passage in the Krishnan patent:

For example, the **bandwidth throttling system** can delay a first class of services provided by a network server in response to the **effective bandwidth utilized by this network server** exceeding a first threshold. If the demand for the bandwidth by this network server exceeds a second threshold, the bandwidth throttling system escalates the throttling response and blocks the first class of services from execution and can also concurrently delay execution of a second class of services. (column 3, lines 1-9; emphasis added)

This passage from the Krishnan patent emphasizes that a server's overload status is not even of concern to Krishnan; rather, Krishnan assumes that the network server is never overloaded and cares only whether or not the *effective network bandwidth* used by the server is greater than a threshold. Accordingly, Krishnan neither discloses nor suggests "determining whether the **server** is **overloaded**" as required by Claim 1.

The passage at column 5, lines 50-65 (quoted in relevant part below) is relied on in the Office Action as supposedly suggesting the claimed step of: "if the **server** is **overloaded** then storing the incoming requests in the buffer only if a number of incoming requests currently stored in the buffer is less than an acceptance limit;" However, neither the Office Action nor the Krishnan patent mentions anything related to whether or not the *server* is overloaded.

Moreover, Krishnan's Asynchronous Thread Queue (ATQ, Krishnan's FIG. 1) is not located within any one network server NS_n (required explicitly by Claim 11), and accordingly it does not make sense for the rejection to attempt to equate the ATQ with the claimed buffer. As shown in Applicants' non-limiting example of a server in FIG. 3, the buffer is associated with a particular, specific processor 300; in contrast, Krishnan's ATQ allocates work among plural network servers.

The passage at column 6, lines 8-12 is relied on in the Office Action as supposedly suggesting the step of: "if the **server** is not **overloaded**, the storing the incoming request in the buffer." However, this passage of the Krishnan patent merely relates to the situation in which a requested operation is stored in the ATQ if the requested operation is "permitted"--*not* based on whether or not a *server* is busy:

...The asynchronous thread queue ATQ queries the bandwidth throttling system BT at step 204 to ensure that the **requested operation is permitted** for the identified network service. This determination is made at step 205 where the bandwidth throttling system BT retrieves the **effective bandwidth measure** for the identified network server NS1 and compares this value with data stored in a control table which is indicative of the **network bandwidth** allocated to this network server NS1. (column 5, lines 53-61; emphasis added)

Again: Krishnan's concern *network* bandwidth does not disclose or suggest any determination of, or even any concern with, whether or not a *server* is overloaded.

Accordingly, Applicants respectfully assert that the Krishnan patent does not disclose or suggest any of the foregoing claim features, especially considered in combination. Accordingly, reconsideration and withdrawal of the rejection of independent Claim 1 are respectfully requested.

Independent Claim 11 recites essentially the same features as Claim 1, discussed above. Accordingly, Claim 11 should be allowable for the same reasons that Claim 1 is allowable. Therefore, reconsideration and withdrawal of the rejection of Claim 11 are respectfully requested.

Dependent Claims 2-10 and 12-20 are allowable for at least the same reasons that Claims 1 and 11 are allowable. Moreover, the dependent claims are allowable based on the merits of their own recitations. For example, Claims 2 and 12 require a determination of a processing rate of the server itself; however, the passage to which the Office Action refers (top of column 3) is merely concerned with an outside force purposely *delaying* the server from providing classes of service based on *network bandwidth* used by a server (regardless of whether the server is overloaded).

Accordingly, reconsideration and withdrawal of the rejection of all claims are respectfully requested.

Rejection under 35 U.S.C. § 103

Claims 3-5 and 13-15 were rejected under 35 U.S.C. § 103 as being unpatentable over Krishnan in view of U.S. Patent No. 6,154,769 (Cherkasova *et al.*). Claims 6(16) and 8-10 (18-20) were rejected under 35 U.S.C. § 103 as being unpatentable over Krishnan in view of U.S. Patent No. 6,510,214 (McDaniel).

Applicants respectfully traverse the rejections.

The Cherkasova *et al.* patent merely relates scheduling requests in a buffer, not to make a decision of whether a server is “overloaded” as claimed. Likewise, McDaniel merely detects overload in a service control point in a communications network, and not on using the claimed “acceptance limit” when a server is overloaded. Accordingly, the Cherkasova *et al.* and McDaniel patents do not overcome the shortcomings of the Krishnan patent in the context of the present invention as discussed above.

Indeed, the Office Action does not even assert that these secondary references disclose or suggest the features discussed above at length in overcoming the rejection under 35 U.S.C. § 102. Applicants' foregoing arguments demonstrate that Claims 1 and 11, and therefore all their dependent claims, are patentably distinguishable over the references, considered either individually or in combination. Accordingly, reconsideration and withdrawal of the rejection under 35 U.S.C. § 103, and allowance of all claims, are respectfully requested.

Request to Verify Change of Correspondence Address; Appointment of New Associate Attorney

A "Notice of Change of Address" has been filed in this application to direct the PTO to send correspondence to a new correspondence address (also set forth below). If the Notice has not been matched with the PTO's application file, the examiner is requested to call the undersigned counsel so that another copy can be provided. If the Notice has been matched with the PTO file, it is requested that the examiner verify that the information has been entered into the PTO mailing system so that future communications will be mailed to the correct address.

Accompanying this Amendment is a "Revocation of Associate Power and Appointment of New Associate Attorney" (combined with the Change of Correspondence Address). The Appointment recognizes the undersigned attorney to prosecute the case.

Conclusion

All objections and rejections have been complied with, properly traversed, or rendered moot. Thus, it now appears that the application is in condition for allowance. Favorable consideration and allowance are earnestly solicited.

Respectfully submitted,

Date: May 21, 2004

For Customer No. 26,652
AT&T Corp.
Room 2A-207
One AT&T Way
Bedminster, New Jersey 07921

By: 

Raymond C. Glenny
Registration No. 32,413
Telephone: 301.464.0431